Energy 80 (2015) 428–436

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

Short-term natural gas demand prediction based on support vector regression with false neighbours filtered

L. Zhu [a], M.S. Li [b], Q.H. Wu [a, *], L. Jiang [a]

[a] Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool L69 3GJ, United Kingdom
[b] School of Electric Power Engineering, South China University of Technology, Guangzhou 510641, China

ARTICLE INFO

Article history:
Received 12 June 2014
Received in revised form
27 November 2014
Accepted 30 November 2014
Available online 19 December 2014

Keywords:
Short-term prediction
Natural gas demand
Time series reconstruction
Support vector regression
Local predictor
False neighbours

ABSTRACT

This paper presents a novel approach, named the SVR (support vector regression) based SVRLP (support vector regression local predictor) with FNF-SVRLP (false neighbours filtered-support vector regression local predictor), to predict short-term natural gas demand. This method integrates the SVR algorithm with the reconstruction properties of a time series, and optimises the original local predictor by removing false neighbours. A unified model, named the SM ("Standard Model"), is presented to process the entire dataset. To further improve the predicted accuracy, an AM ("Advanced Model") is proposed, and is based on specific customer behaviours during different days of the week. The AM contains seven individual models for the seven days of the week. The FNF-SVRLP based AM has been used to predict natural gas demand for the National Grid of the United Kingdom (UK). This model outperforms the SVRLP, the ARMA (autoregressive moving average) and the ANN (artificial neural network) methods when applied to real-world data obtained from National Grid and has been successfully applied to daily gas operations for National Grid.

1. Introduction

Energy is a vital element in the development of societies and economies worldwide [1]. Energy modelling and forecasting has attracted increased attention within energy sectors because of environmental concerns and increasingly stringent government policies on energy generation. Moreover, the absence of a reliable and uninterrupted energy supply can be a bottleneck for the economical and societal development of a country; to provide a framework for the development of a country's future energy supplies, it is therefore important for policymakers to understand future energy demands [2]. Natural gas as a primary energy source, was first commercially used in approximately 1785, and has since been used widely in Europe [3,4]. The UK has become the largest market in Europe over the past decade. According to HI-Energy [5] and government reports from the UK [6], 46% of the total electricity (approximately 350,000 GWh) generated and approximately 70% of domestic heating in the UK in 2008 is produced by natural gas. Natural gas plays a key role in the UK's energy mix. However, natural gas is a non-renewable energy source, has limited reserves on the planet and may contribute to global warming. Thus, planning for future energy demand is becoming an important issue in the energy sector, and such planning must begin with the forecasting of such demand, particularly for natural gas. Furthermore, several countries (for example, the UK) are highly dependent on the import of gas. Therefore, an accurate gas demand prediction model can directly lower the purchasing costs for distributors and end consumers. Therefore, accurate gas demand prediction models could increase the efficiency of natural gas usage with an optimal purchasing cost and lower a gas that may contribute to global warming.

Studies in the field of NGDP (natural gas demand prediction) can be classified as long-term NGDP and short-term NGDP. Long-term predictions focus on forecasting demand over a long time period, i.e., on an annual or monthly basis. In 2003, Sarak and Satman [7] applied a degree-day method to forecast the potential natural gas demand in Turkey from 2004 to 2023. Imam et al. [8] presented the multicyclic Hubbert model to estimate future natural gas demand consumption trends. Gutierrez et al. [9] used a Gompertz-type innovation diffusion process as a stochastic growth model for predicting natural gas demand. Potočnik et al. [10] presented a statistics-based machine forecasting model to predict the future natural gas demand in Slovenia in 2005 and 2006. Erododu [11]

* Corresponding author.
E-mail address: qhwu@liv.ac.uk (Q.H. Wu).

http://dx.doi.org/10.1016/j.energy.2014.11.083
0360-5442/

applied an autoregressive integrated moving average model to forecast the future growth in gas demand in Turkey.

An accurate short-term NGDP is equally important as long-term predictions because distributors are required by their suppliers to provide the amount of natural gas they require for the coming hours or days within a regulated tolerance interval. Most short-term NGDP methods employ ANN (artificial neural network). Brown et al. [12] and Brown and Matin [13] used an ANN to predict natural gas demands, which demonstrated a higher level of accuracy compared to a linear regression model. Khotanzad and Elragal [14] and Khotanzad et al. [15] combined an ANN with various methods and proposed a two-stage prediction models, which demonstrated that different combinations of ANN models can improve the prediction accuracy. In Serbia, Ivezić [16] used the daily gas demand and daily minimal and maximal temperatures as inputs to perform a short-term NGDP with an ANN. Azadeh et al. [17] applied an ANN-based algorithm, named ANFIS (adaptive neuro-fuzzy inference system), to predict the natural gas demand in the Iranian network. They applied the gas demand from the same day in the previous year as an additional input with other conventional inputs.

A SVM (support vector machine) is a powerful machine learning method that is based on statistical learning theory [18]. The ERM (empirical risk minimisation) principle, which is generally employed in traditional ANN, is replaced by SRM (structural risk minimisation) principle in a SVM. The most important concept of SRM is the utilisation of minimising an upper bound to generalise the error instead of minimising the training error. With the introduction of Vapniks ε-insensitive loss function, SVMs have been extended to solve regression problems called SVR (support vector regression) [19]. Recently, SVR has been applied to various problems with promising results [20,21].

Because of the complexity of historical gas demand and the uncertainty behind the factors influencing the demand, such as meteorological, economical, and other random factors, the time series reconstruction technique can be applied to gas demand forecasting. According to previous studies [22], local prediction methods based on phase reconstruction typically perform better than global methods based on phase reconstruction. With the local prediction method, each predicting point has its own model constructed based on its NNs (nearest neighbours), which are found in the neighbourhood of the phase space reconstructed from the time series, and the fitness of the NNs would mainly affect the model performance. However, NNs may contain a class of FNs (false neighbours), which would dramatically decrease the fitting accuracy and result in a poor modelling performance. Therefore, not all NNs are suitable for local predictions, and several NNs should be filtered [23].

In the present work, a false neighbours filter to be combined with the SVR and space reconstruction of a time series is proposed. The resulting predictor is referred to as SVRLP (support vector regression local predictor) with FNs filtered (FNF-SVRLP (false neighbours filtered-support vector regression local predictor)) and is examined for its applicability to short-term gas demand forecasting. FNF-SVRLP is a modified SVRLP, that improves the LP (local predictor) by optimising the NNs. These NNs play important roles in improving the accuracy of local modelling. The FNs filtering algorithm uses a combination of an exponential separation rate and Euclidian distance to optimise the NNs. We propose a unified model, named the SM ("Standard Model"), that evenly treats the entire dataset. The customer behaviour based AM ("Advanced Model") is then presented. The AM is a combination of seven models, which are Mon. Model, Tue. Model, Wed. Model, Thu. Model, Fri. Model, Sat. Model, and Sun. Model. The proposed FNF-SVRLP is verified using an operational gas dataset provided by National Grid and is compared with the ARMA (autoregressive moving average), ANN and SVRLP to demonstrate its superior performance.

2. Time series reconstruction

Packard [24] noted that the phase space can be reconstructed from a univariate time series, because of the nature of the time series information of all of the variables in this dynamic system. Takens [25] and Sauer [26] developed the embedding theorem, which provided support Packard's concept. The theorem treats a one-dimensional time series as compressed information of a higher dimension. An univariate one-dimensional time series $x(t)$ for $t = 1, 2, 3 \ldots T$ can be extracted by extending $x(t)$ to a vector $X(t)$ in a m-dimensional space as follows:

$$X(t) = [x(t),\ x(t-\tau),\ x(t-2\tau), \ldots,\ x(t-(m-1)\tau)] \tag{1}$$

where m is the embedding dimension and τ is the delay constant. After reconstructing the one-dimensional time series to a higher dimensional time series, the obtained time series is called a reconstructed time series. Numerous of methods calculate the attractor and delay constant. For example, the correlation dimension method [27] is the most popular method for determining m, and the mutual information method [28] is a suitable method for selecting of τ. The above methods cannot be easily applied to multivariate time series. Therefore, we present an accuracy-based method to estimate the embedding dimension and delay time. This step is discussed in Section 5.2.1.

3. Support vector regression based local predictor with false neighbours filtered

3.1. Support vector regression

The purpose of a regression is to approximate a function, $g(x)$, from a given dataset, $G = (x_i, y_i)_{i=1}^{N}$, obtained from g with noise. The main concept of SVR is to map the input data x into a high dimension feature space via a nonlinear mapping and subsequently perform a linear regression to estimate the following function in terms of the new feature space.

$$f(x) = (w, x) + b, \tag{2}$$

where $(\cdot,\cdot)$ is the dot product on the input space, w is the function coefficient vector, and b is a real constant called the "threshold value". To estimate the function f, a loss function called Vapnik's ε-insensitive function is given as follows:

$$|y - f(x)|_{\varepsilon} = \begin{cases} 0 & \text{if} |y - f(x)| \le \varepsilon, \\ |y - f(x)| - \varepsilon & \text{otherwise}, \end{cases} \tag{3}$$

where ε is the tolerance to error, and only those deviations larger than ε are considered to be errors. The coefficients w and b can be determined from the data by minimising the following function

$$R(w) = \frac{1}{2}||w||^2 + \frac{C}{N}\sum_{i=1}^{N}|y_i - f(x_i)|_{\varepsilon} \tag{4}$$

where C is regularisation constant, that determines the trade-offs between the training errors and model complexity. This problem can be transformed into a constrained optimisation problem by introducing the positive slack variables ξ and ξ^* as follows:

$$\begin{aligned}&\min\ R(w)=\frac{1}{2}||w||^2+\frac{C}{N}\sum_{i=1}^{N}(\xi_i+\xi_i^*),\\&\text{s.t.}\\&f(x_i)-y_i\le\varepsilon+\xi_i,\\&y_i-f(x_i)\le\varepsilon+\xi_i^*,\\&\xi_i,\xi_i^*\ge 0\quad\text{for } i=1,2,\ldots,N.\end{aligned}\tag{5}$$

The slack variables ξ and ξ^* can be introduced when the data cannot be estimated by the function f under the precise ε.

Lagrange multipliers can be added to the equations to solve the optimisation problem. The above problem can then be written as its dual form:

$$\begin{aligned}\max L_p=&-\varepsilon\sum_{i=1}^{N}(\alpha_i+\alpha_i^*)+\sum_{i=1}^{N}(\alpha_i-\alpha_i^*)y_i-\frac{1}{2}\sum_{i,j=1}^{N}(\alpha_i-\alpha_i^*)\left(\alpha_j-\alpha_j^*\right)\\&\times(x_i,x_j),\text{s.t. } 0\le\alpha_i,\alpha_i^*\le C\text{ and }\sum_{i=1}^{N}(\alpha_i-\alpha_i^*)=0,\end{aligned}\tag{6}$$

where α_i and α_i^* are the Lagrange multipliers. For nonlinear cases, the dot product (x_i, y_i) shown in Equation (3) may be replaced by a kernel $\Phi(x_i\cdot\Phi(x_j))=k(x_i,x_j)$, which is defined as a Gaussian kernel function as follows:

$$k(x,x')=\exp\left(-\frac{||x-x'||^2}{2\alpha^2}\right),\tag{7}$$

Then, Equation (2) can be written as follows:

$$f(x)=\sum_{i=1}^{N}(\alpha_i-\alpha_i^*)k(x_i,x)+b.\tag{8}$$

3.2. Local predictor with FNs filtered

The local predictor relies on a set of nearest neighbours that evolves similarly in the reconstructed phase space. The choice of neighbours is affected and limited by the finite size of the data set, by the stochastic noise, and by the complex structure of the attractor. This limitation is the main source of errors in the analysis. Finding suitable and reasonable neighbours of one known point is one of the most important tasks in achieving reliable results. The original method of selecting the subset (neighbourhoods) is based on the Euclidian distance between the testing data and training data in the input space. For each query vector q, the K nearest neighbours $\{g_q^1, g_q^2, \ldots, g_q^K\}$ among the training inputs are selected using the Euclidian distance as the distance metric between q and each g in the reconstructed time series.

$$\left|\left|q-g_q^K\right|\right|>\left|\left|q-g_q^{K-1}\right|\right|$$

However, FNs could exist in the reconstructed state space after using a particular embedding method. Such FNs can be caused by an improper embedding dimension, such as one that is not sufficiently large. This effect is illustrated in Fig. 1. Suppose that the predicting (reference) point is $X_n(t)$. By utilising the original local predictor (Euclidean distance), two NNs are found, $X_{n+1}(t)$ and $X_{n+2}(t)$, and $||X_n(t)-X_{n+2}(t)||<||X_n(t)-X_{n+1}(t)||$. When only considering of the Euclidean distance, $X_{n+2}(t)$ is a true neighbour of $X_n(t)$. However, because of the large deviation between $X_{n+2}(t+1)$ and the predicting (target) point $X_n(t+1)$, the selection of $X_{n+2}(t)$ as a training sample would dramatically decrease the fitting accuracy. Therefore, $X_{n+1}(t)$ is superior to $X_{n+2}(t)$, even when $X_{n+2}(t)$ is closer to the reference point. Neighbours as $X_{n+2}(t)$ are called FNs. Selected these points as the training samples, would decrease the accuracy of the local modelling. To filter FNs, we must determine the validity of each neighbour. In this paper, we present a convenient method that uses the exponential separation rate to verify the neighbours. In the reconstructed phase space, $X(t)$ is set as the predicting (reference) point at time t, and $X_j(t)$ is the jth neighbour of $X(t)$ in the phase space, where $1\le j\le N$. The exponential separation rate between the track of $X(t)$ and the track of $X_j(t)$ is as follows:

$$\xi_j=\ln\left|\frac{d_j(t)}{d_j(t-1)}\right|\tag{9}$$

where $d_j(t-1)$ is the Euclidean distance between $X(t-1)$ and $X_j(t-1)$ at time $t-1$. The proposed method employs the Euclidean distance d_j and exponential separation rate ξ_j to determine the validity of $X_j(t)$, which is $1\le j\le N$ and $\xi_j\le\varphi$. This method ensures that $X_j(t)$ is one of the nearest neighbours with a small exponential separation rate, and the values of N and φ can be modified (increased or decreased) to obtain K optimal nearest neighbours.

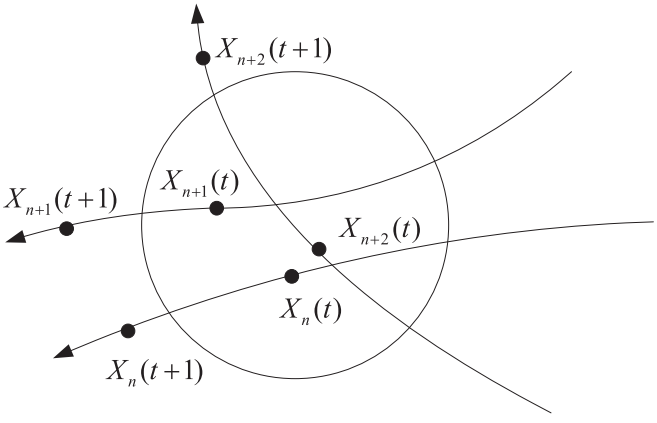

Fig. 1. Schematic diagram of the effect of the forecast on the evolutionary track of neighbouring points. In this figure, $X_{n+2}(t)$ is the FN point of $X_n(t)$.

4. Implementation of the proposed FNF-SVRLP

4.1. SVRLP

The SVRLP method is a general SVR-based local predictor [22] that has been previously presented. The SVRLP method can be summarised as follows: First, τ and m are used to reconstruct the time series. Then, the local predictor (Euclidean distance) is applied to determine the K nearest neighbours (which could contain FNs) for each query vector. The K nearest neighbours are then used to train the SVR to obtain support vectors and the corresponding coefficients. Finally, the output can be computed by Equation (8).

4.2. FNF-SVRLP

From Section 3.2, FNs can exist in the reconstructed state space; selecting such FNs as the training samples, would decrease the accuracy of the local modelling. Therefore, NNs with FNs must be filtered before being used to train the SVR. The steps for gas demand prediction based on the proposed method can be summarised as follows:

- Step 1 Reconstruct the time series: Load the multivariate time series dataset $X = (x_1(t), x_2(t), \ldots, x_i(t))$, $(t = 1, 2, \ldots, T)$. Use the accuracy-based method to determine the embedding dimension m and time delay constant τ. Then, reconstruct the multivariate time series using the embedding dimension m and time delay constant τ, additional details are presented in Section 5.2.1.
- Step 2 Form a training and validation data: The input dataset after reconstruction is denoted as $\tilde{X}$ and the size of $\tilde{X}$ is N_{all}. Then, divided the $\tilde{X}$ into two parts: a training $\tilde{X}_{tr}$ and a validation $\tilde{X}_{va}$. The size of the training dataset is N_{tr}, the size of the validation dataset is N_{va}, and $N_{all} = N_{tr} + N_{va}$.
- Step 3 Determine the nearest neighbours: Set a query point at time t as $x_q(t)$ and a point in $\tilde{X}_{tr}$ as $\tilde{X}_{tr_i}$ where $\tilde{X}_{tr} = (\tilde{X}_{tr_1}, \tilde{X}_{tr_2}, \ldots, \tilde{X}_{tr_i})$, $(i = 1, 2, \ldots, N_{tr})$. Then, calculate the Euclidian distance D_i $(i = 1, 2, \ldots, N_{tr})$ between the $x_q(t)$ and $\tilde{X}_{tr_i}$, where $D_i = \sqrt{(\tilde{X}_{tr_i} - x_q(t))^2} (i = 1, 2, \ldots,)$. Then, re-rank D_i from smallest to the largest, and the N corresponding nearest neighbours are denoted as $\{z^1_{x(t)_q}, z^2_{x(t)_q}, z^3_{x(t)_q}, \ldots z^N_{x(t)_q}\}$ $(1 < N \ll N_{tr})$
- Step 4 Calculate the exponential separation rate: For the N nearest neighbours $\{z^1_{x(t)_q}, z^2_{x(t)_q}, z^3_{x(t)_q}, \ldots, z^N_{x(t)_q}\}$, each nearest neighbour can obtain its correspondent exponential separation rate via Equation (9), which is $\{\xi_{z^1_{x(t)q}}, \xi_{z^2_{x(t)q}}, \xi_{z^3_{x(t)q}}, \ldots, \xi_{z^N_{x(t)q}}\}$. Additional details are given in Section 3.2.
- Step 5 Obtain K optimal nearest neighbours: First, the Equation (12) is applied to calculate the value of K. The ith optimal nearest neighbour must meet the following requirements:(a) $1 \le i \le K < N$; (b) $\xi_{z^i_{x(t)_q}} < \varphi$. This ensures that the optimal nearest neighbour is a nearest neighbour with a small exponential separation rate. The values of N and φ are modified until K optimal nearest neighbours are obtained. Additional details are presented in Section 3.2.
- Step 6 Train the SVR: Use the K optimal nearest neighbours of the query point to train the SVR algorithm.
- Step 7 Calculate the prediction value of the current query point using Equation (8).
- Step 8 Then, repeat steps 3 to 7 until the future values of different query points are all acquired.

Fig. 2 presents the computation procedure for the proposed gas demand prediction method.

5. Case study

Case studies are based on the recorded gas data from National Grid from the UK. First, the FNF-SVRLP based SM (Section 5.3) is used to process the entire gas dataset. Then, based on the results obtained from the SM, the accuracy of the proposed SM could be improved by considering the customer behaviour difference within one week. We upgraded the SM to the AM (Section 5.4), which includes seven models (the Mon. Model, Tue. Model, Wed. Model, Thu. Model, Fri. Model, Sat. Model, and Sun. Model). The details are described in the following.

5.1. Data description

The dataset used in the research was provided by National Grid. There is a data item explorer on National Grid's website [29], which

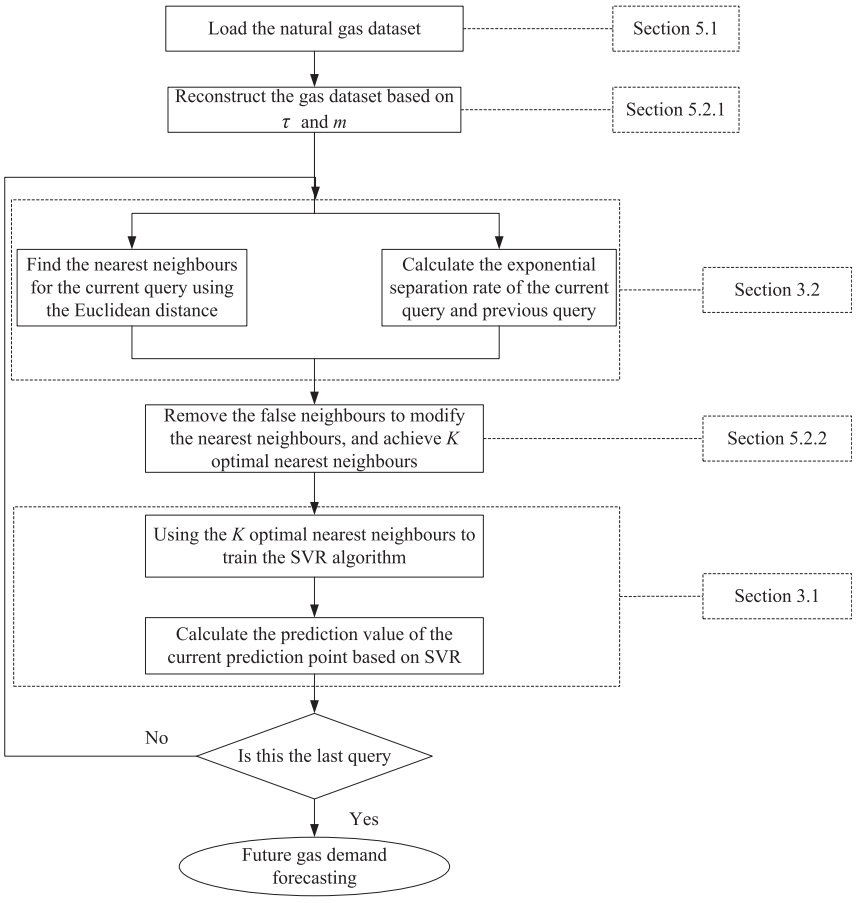

Fig. 2. Flowchart of the proposed model.

Fig. 3. UK local distribution zones.

provides market participants and shippers with information related to gas. After a number of tests, the following datasets obtained from the data item explorer were selected as our experimental data:

- NTS: The NTS (National Transmission System) demand refers to the amount of gas used by gas consumers directly connected to the NTS and all local distribution zones (13 local distribution zones operate in the UK, as shown in Fig. 3).
- CWV: The CWV (Composite Weather Variable) is a single measure of weather for each local distribution zone that accounts, for not only temperature, but also wind speed, effective temperature and the pseudo seasonal normal effective temperature.

The NTS gas demand and 13 CWVs are recorded at one day interval. The time duration of the dataset is from Jan. 1st, 2009 to Dec. 31st, 2012 (Dataset A in Table 1), or four years. Dataset A is then divided into two subsets, Dataset B_1 and Dataset B_2. We utilise Dataset B_1 which contains the first three years (from Jan. 1st, 2009 to Dec. 31st, 2011) to develop a prediction model for the NTS gas demand. Dataset B_2 (from Jan. 1st, 2012 to Dec. 31st, 2012) is used to test the prediction performance of the model learned from Dataset B_1.

For all performed models, we quantified the prediction performance with the MAPE (mean absolute percentage error) and the MAE (mean absolute error). These parameters can be defined as follows:

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\frac{|\hat{y}_i - y_i|}{y_i} \times 100 \tag{10}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i| \tag{11}$$

where N is the size of the testing dataset, $\hat{y}_i$ is the forecasted gas demand, y_i is the actual gas demand, $\bar{y}$ is the mean of the actual gas demand, and i is the test instance index.

5.2. Setting the parameters

5.2.1. Embedding dimension and delay constant

In this paper, the accuracy-based method is introduced to estimate the embedding dimension m and the delay time τ. We denote day i as D_i, the previous day as D_{i-1} and the next day as D_{i+1}. To forecast the NTS gas demand G_{i+1} of D_{i+1}, we summarise the variables known a priori in Table 2.

In Table 2, G_i is the NTS gas demand of day D_i, $C_i(13)$ are the actual CWVs of the 13 local distribution zones on day D_i and $C^f_{i+1}(13)$ are the forecasted CWVs of the 13 local distribution zones on day D_{i+1}. The above data can be obtained from National Grid's website. We then apply the SVRLP based SM model to Dataset B_1 to help us estimate the embedding dimension and delay constant. The results are presented in Tables 3 and 4. The threshold MAPE value of 3.5% obtained in the computation has been considered to be a high-quality result. A lower threshold value produces more predictors. A large number of predictors can result in extracted models that exhibit inferior performance because of "the curse of dimensionality" principle [30]. Therefore, the embedding dimension m is 55 and the delay time τ is 1.

5.2.2. K nearest neighbours

The selection of K is important in establishing the local predictor. Several methods used in the literature can be used to determine this parameter, such as the cross validation [31] and bootstrap [32] methods. This parameter should be small for high density datasets and large for low density datasets. In this paper, we calculate K using a previously proposed systematic method [33] as follows:

$$K = \mathrm{round}\left(\frac{\partial}{H \times k_{\max} \times D_{\max}}\sum_{i=1}^{H}\sum_{k=1}^{k_{\max}} D_k(x_i)\right) \tag{12}$$

where H is the total number of training points, $k_{\max}$ is the maximum number of nearest neighbours, $D_k(x_i)$ is the distance between each training point x and its nearest neighbours, $D_{\max}$ is the maximum distance, $\frac{1}{H\times k_{\max}\times D_{\max}}\sum_{i=1}^{H}\sum_{k=1}^{k_{\max}} D_k(x_i)$ is the average

Table 1
Dataset description.

Data set	Start time stamp	End time stamp	Description
A	01/01/2009	31/12/2012	Total dataset; 1461 days
B_1	01/01/2009	31/12/2011	Train dataset; 1095 days
B_2	01/01/2012	31/12/2012	Test dataset; 366 days

Table 2
Obtained variables.

	Target	Prior-known			
Day	D_{i+1}	D_i	D_{i-1}	D_{i-2}	… D_1
NTS gas demand	G_{i+1}	G_i	G_{i-1}	G_{i-2}	… G_1
Actual CWVs	–	$C_i(13)$	$C_{i-1}(13)$	$C_{i-2}(13)$	… $C_1(13)$
Forecast CWVs	–	$C^f_{i+1}(13)$	–	–	… –

Table 3
MAPE for different values of embedding dimension with $\tau = 1$.

M	m	τ=1				MAPE
1	27	D_i				4.1%
2	41	D_i	D_{i-1}			3.8%
3	55	D_i	D_{i-1}	D_{i-2}		3.5%
4	69	D_i	D_{i-1}	D_{i-2}	D_{i-3}	3.4%

Table 4
MAPE for different values of embedding dimension with τ=2

M	m	$\tau = 2$				MAPE
1	14	D_{i-1}				5.4%
2	28	D_{i-1}	D_{i-3}			5.1%
3	42	D_{i-1}	D_{i-3}	D_{i-5}		4.7%
4	56	D_{i-1}	D_{i-3}	D_{i-5}	D_{i-7}	4.3%

distance around the points, which is inversely proportional to the local densities and ∂ is a constant. Constants k_{max} and ∂ are parameters with low sensitivity. k_{max} can be selected as a percentage of the number of training points (H) to improve efficiency, where ∂ can be selected as a percentage. In this paper, k_{max} and ∂ are always fixed after a few trials for all test cases at 35% of H and 150, respectively.

5.3. Standard model

The SM is a unified model, and is applied to process the entire dataset. We denote the week j as W_j, the previous week as W_{j-1} and the following week as W_{j+1}. The input dataset for the SM is shown in Table 5. A comparison of the SM with ARMA, ANN, SVRLP and FNF-SVRLP is presented in Section 5.3.1.

5.3.1. Results

To evaluate the performance of the proposed FNF-SVRLP, comparisons with ARMA, ANN and SVRLP are conducted. The comparison results are shown in Table 6 and Fig. 4. The results indicate that, SVRLP performs better than ARMA and ANN, FNF-SVRLP is the best among all four methods. The MAPE of FNF-SVRLP is 3.8%, which is an improvement over those of the ARMA, ANN and SVRLP methods by 60.5%, 39.4% and 15.7%, respectively. Additionally, the MAE of the FNF-SVRLP is 9.2 mcm (million cubic meters), which is smaller than the other MAE values. Using the MAE of SVRLP subtracted by the MAE of FNF-SVRLP, the result we achieved was approximately 1.2 mcm, which indicates that the FNF-SVRLP could save an additional 1.2 mcm of gas in terms of demand per day compared to SVRLP. Figs. 5 and 6 show the MAPE and MAE for every day of the week (Monday-Sunday) during the testing period. These results confirm the superiority of the FNF-SVRLP over the other three methods. In addition, the MAPE and MAE of the entire testing data over 12 months are calculated and shown in Table 7. The performance of FNF-SVRLP is better than that of ARMA, ANN and SVRLP. Figs. 7–10 show the results of the natural gas demand prediction by different methods in comparison of actual natural gas demand. The aggregated distribution of MAPEs' for FNF-SVRLP and SVRLP is illustrated in Fig. 11. The above results clearly indicate the accuracy of the SVR-based local predictor and the importance of filtering FNs. Filtering FNs while maintaining optimal neighbours and using those optimal neighbours as the training samples improve the accuracy.

5.4. Advanced model

According to the results obtained from the SM (Section 5.3.1), the accuracy for Saturday and Monday are always lower than for the other days of a week (see Figs. 5 and 6). In the UK, Monday is the start of the weekday, and Saturday is the start of the weekend; therefore, the customer behaviour for these two days will be very different from the other days. However, customer behaviour is an immeasurable factor. To eliminate the effect of this behaviour, we presented the Advanced Model, which includes seven individual models for each day of the week, as shown in Tables 8–14. Comparisons between SVRLP based AM and FNF-SVRLP based AM are presented in Section 5.4.1.

Table 5
Input dataset for SM.

Target	Input		
.	.	.	.
Mon.(W_{j+1})	Sun.(W_j)	Sat.(W_j)	Fri.(W_j)
Sun.(W_j)	Sat.(W_j)	Fri.(W_j)	Thu.(W_j)
Sat.(W_j)	Fri.(W_j)	Thu.(W_j)	Wed.(W_j)
Fri.(W_j)	Thu.(W_j)	Wed.(W_j)	Tue.(W_j)
Thu.(W_j)	Wed.(W_j)	Tue.(W_j)	Mon.(W_j)
Wed.(W_j)	Tue.(W_j)	Mon.(W_j)	Sun.(W_{j-1})
Tue.(W_j)	Mon.(W_j)	Sun.(W_{j-1})	Sat.(W_{j-1})
Mon.(W_j)	Sun.(W_{j-1})	Sat.(W_{j-1})	Fri.(W_{j-1})
Sun.(W_{j-1})	Sat.(W_{j-1})	Fri.(W_{j-1})	Thu.(W_{j-1})
.	.	.	.

Table 6
MAPE and MAE of different methods using the SM.

	MAPE	MAPE impro.	MAE	MAE impro.
FNF-SVRLP	3.8%	–	9.2	–
SVRLP	4.4%	15.7%	10.4	13.0%
ANN	5.3%	39.4%	11.9	29.3%
ARMA	6.1%	60.5%	13.6	47.8%

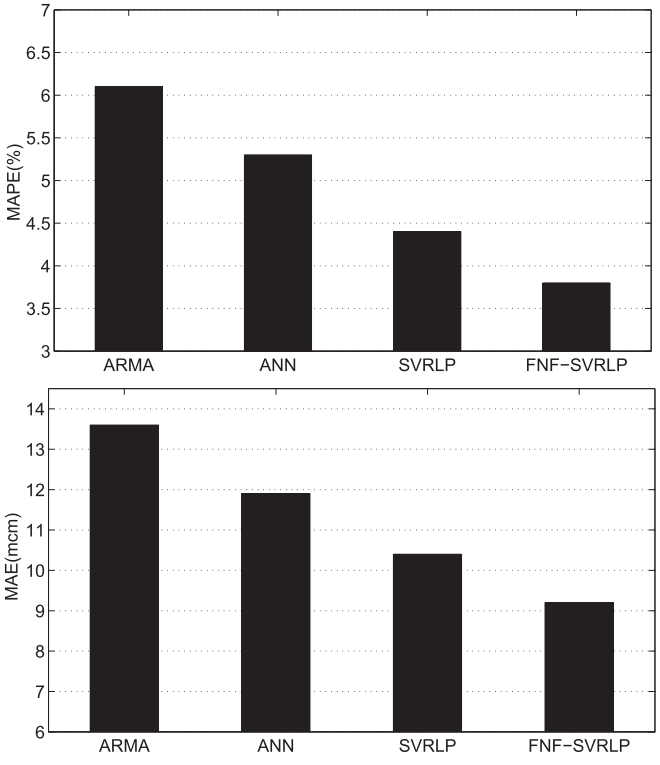

Fig. 4. Comparison of the FNF-SVRLP method and the other methods.

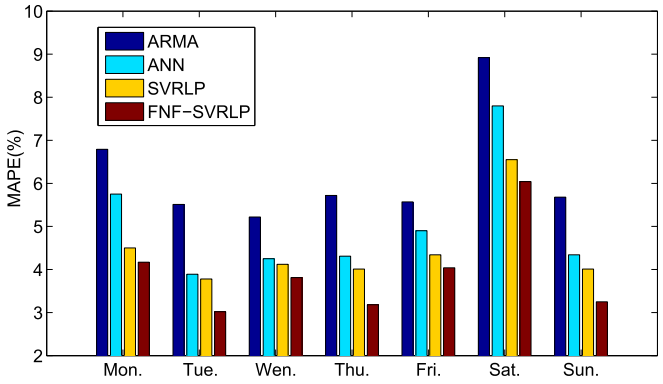

Fig. 5. Average prediction MAPE for every day of the week during the entire testing period.

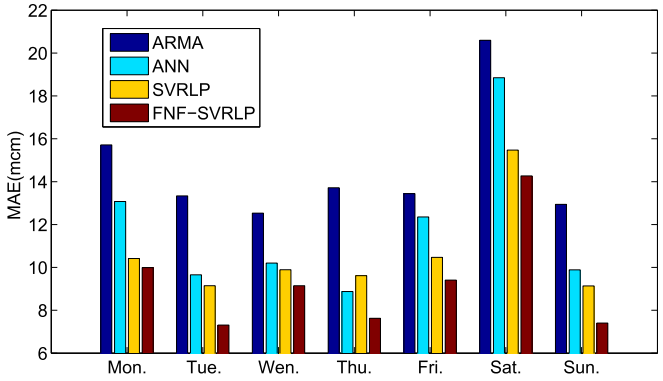

Fig. 6. Average prediction MAE for every day of the week during the entire testing period.

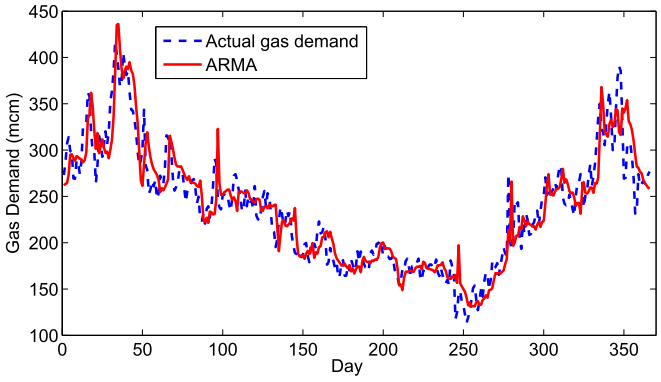

Fig. 7. The comparison of the predicted gas demand calculated by the ARMA method and the actual gas demand for the entire testing period.

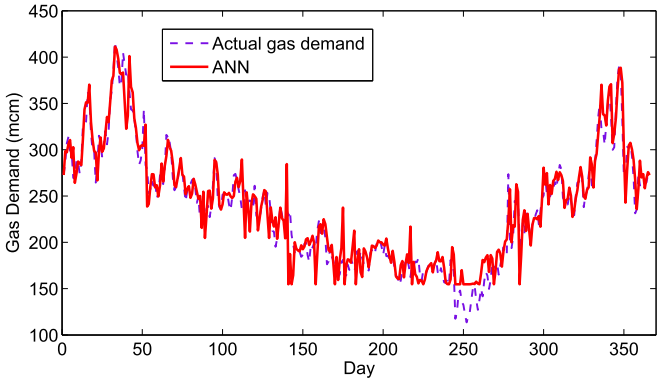

Fig. 8. The comparison of the predicted gas demand calculated by the ANN method and the actual gas demand for the entire testing period.

5.4.1. Results

The results obtained from the SM show that the FNF-SVRLP provides significantly better forecast results compared to ARMA, ANN and SVRLP. However, the MAPE of Monday and the MAPE of Saturday are the top two highest MAPEs among the seven days; an identical pattern is noted for the MAE as well. Considering that the customer behaviour of each day in a week would be different from the others, we presented an AM for these seven individual models to overcome this problem. Table 15 contains the prediction results of the SM and the AM using two different methods. This table shows that regardless of the application of the FNF-SVRLP method and the AM outperforms the other combinations. From Figs. 12 and 13, some improvement is noted in the accuracy over seven days, particularly on Monday, Saturday and Sunday. From the above results, it can be observed that the AM performs better than the SM, which indicates that customer behaviour can affect the forecasting performance.

Table 7
Average prediction MAPE and MAE for every month during the entire testing period.

Month	MAPE (%)				MAE (mcm)			
	ARMA	ANN	SVRLP	FNF-SVRLP	ARMA	ANN	SVRLP	FNF-SVRLP
1	4.6	3.8	3.2	3.0	13.2	11.4	9.8	9.2
2	7.2	5.5	5.1	4.3	20.8	18.2	17.0	14.8
3	6.1	4.2	2.9	2.5	14.1	10.6	7.6	6.5
4	6.7	5.8	4.2	3.9	16.7	14.2	10.4	9.6
5	7.5	5.3	3.6	3.1	12.3	9.5	7.8	6.8
6	6.5	5.5	4.9	3.7	13.2	10.2	8.9	6.7
7	5.1	4.8	4.0	3.6	9.2	8.6	7.2	6.4
8	4.9	5.2	3.9	3.7	8.4	9.3	6.7	6.3
9	11.7	11.5	8.8	7.1	14.6	14.4	12.2	9.8
10	9.4	8.7	7.5	6.6	20.1	18.2	16.9	14.8
11	5.1	3.2	2.6	2.4	11.6	8.5	7.1	6.4
12	6.1	5.1	4.8	4.5	16.4	14.7	14.3	13.4

6. Conclusion

In this paper, we have presented a FNF-SVRLP based AM for short-term gas demand prediction. The forecasted gas demand can aid the market in developing efficient decisions by balancing the

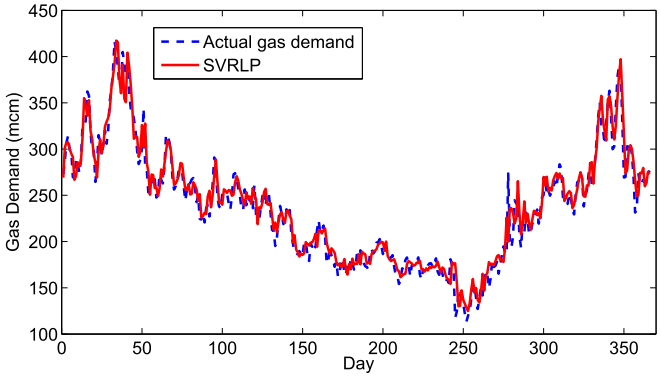

Fig. 9. The comparison of the predicted gas demand calculated by the SVRLP method and the actual gas demand for the entire testing period.

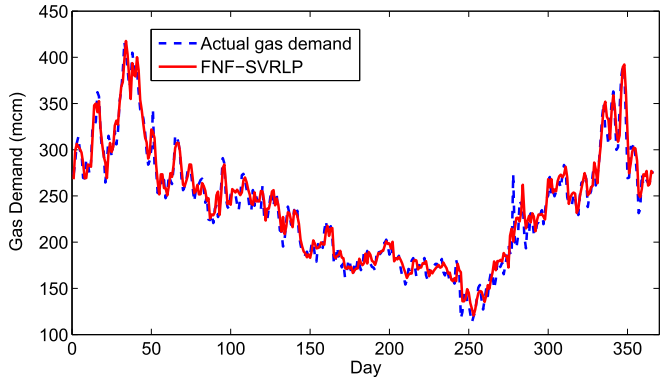

Fig. 10. The comparison of the predicted gas demand calculated by the FNF-SVRLP method and the actual gas demand for the entire testing period.

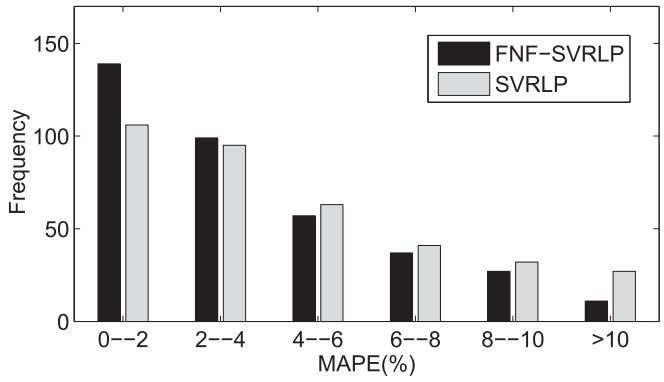

Fig. 11. The aggregated distribution of MAPEs' for FNF-SVRLP and SVRLP during the whole testing period.

Table 8
Input dataset for Monday Model.

Target	Input		
.	.	.	.
Mon.(W_{j+1})	Sun.(W_j)	Sat.(W_j)	Fri.(W_j)
Mon.(W_j)	Sun.(W_{j-1})	Sat.(W_{j-1})	Fri.(W_{j-1})
Mon.(W_{j-1})	Sun.(W_{j-2})	Sat.(W_{j-2})	Fri.(W_{j-2})
.	.	.	.

Table 9
Input dataset for Tuesday Model.

Target	Input		
.	.	.	.
Tue.(W_{j+1})	Mon.(W_{j+1})	Sun.(W_j)	Sat.(W_j)
Tue.(W_j)	Mon.(W_j)	Sun.(W_{j-1})	Sat.(W_{j-1})
Tue.(W_{j-1})	Mon.(W_{j-1})	Sun.(W_{j-2})	Sat.(W_{j-2})
.	.	.	.

Table 10
Input dataset for Wednesday Model.

Target	Input		
.	.	.	.
Wed.(W_{j+1})	Tue.(W_{j+1})	Mon.(W_{j+1})	Sun.(W_j)
Wed.(W_j)	Tue.(W_j)	Mon.(W_j)	Sun.(W_{j-1})
Wed.(W_{j-1})	Tue.(W_{j-1})	Mon.(W_{j-1})	Sun.(W_{j-2})
.	.	.	.

Table 11
Input dataset for Thursday Model.

Target	Input		
.	.	.	.
Thu.(W_{j+1})	Wed.(W_{j+1})	Tue.(W_{j+1})	Mon.(W_{j+1})
Thu.(W_j)	Wed.(W_j)	Tue.(W_j)	Mon.(W_j)
Thu.(W_{j-1})	Wed.(W_{j-1})	Tue.(W_{j-1})	Mon.(W_{j-1})
.	.	.	.

Table 12
Input dataset for Friday Model.

Target	Input		
.	.	.	.
Fri.(W_{j+1})	Thu.(W_{j+1})	Wed.(W_{j+1})	Tue.(W_{j+1})
Fri.(W_j)	Thu.(W_j)	Wed.(W_j)	Tue.(W_j)
Fri.(W_{j-1})	Thu.(W_{j-1})	Wed.(W_{j-1})	Tue.(W_{j-1})
.	.	.	.

Table 13
Input dataset for Saturday Model.

Target	Input		
.	.	.	.
Sat.(W_{j+1})	Fri.(W_{j+1})	Thu.(W_{j+1})	Wed.(W_{j+1})
Sat.(W_j)	Fri.(W_j)	Thu.(W_j)	Wed.(W_j)
Sat.(W_{j-1})	Fri.(W_{j-1})	Thu.(W_{j-1})	Wed.(W_{j-1})
.	.	.	.

Table 14
Input dataset for Sunday Model.

Target	Input		
.	.	.	.
Sun.(W_{j+1})	Sat.(W_{j+1})	Fri.(W_{j+1})	Thu.(W_{j+1})
Sun.(W_j)	Sat.(W_j)	Fri.(W_j)	Thu.(W_j)
Sun.(W_{j-1})	Sat.(W_{j-1})	Fri.(W_{j-1})	Thu.(W_{j-1})
.	.	.	.

supply and demand, thereby reducing costs. The FNF-SVRLP combines the SVR-based local predictor with a FNs filter. During the computation procedure of the gas demand forecasting, the embedding dimension and time delay are computed through an accuracy-based method. Phase space reconstruction is then applied to process the gas dataset. In the local predictor, several differences between the SVRLP and proposed FNF-SVRLP are noted. Compared with the SVRLP, the FNF-SVRLP applies not only the Euclidean distance but also the exponential separation rate to verify the nearest neighbours. Only the optimal nearest neighbours can be selected as training samples. The final results demonstrate that the FNF-SVRLP can achieve a higher prediction accuracy than that of the ARMA method, ANN method and SVRLP method using identical real-world gas data. Moreover, a study on the customer behaviour for different days has been performed, and the results indicate that customer behaviour can affect the final forecasting performance.

Table 15
Prediction results of Standard Model and Advanced Model with different methods under the same testing gas dataset.

	MAPE (%)		MAE (mcm)	
	SVRLP	FNF-SVRLP	SVRLP	FNF-SVRLP
Standard model	4.4	3.8	10.4	9.2
Advanced model	4.0	3.4	9.4	8.3

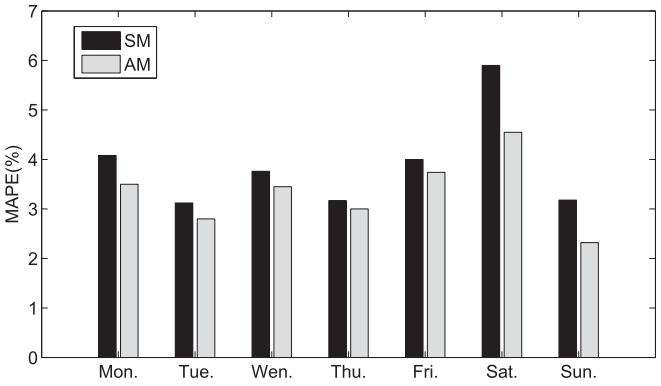

Fig. 12. Comparison of the average prediction MAPE for every day of the week between the SM and the AM by using the FNF-SVRLP method.

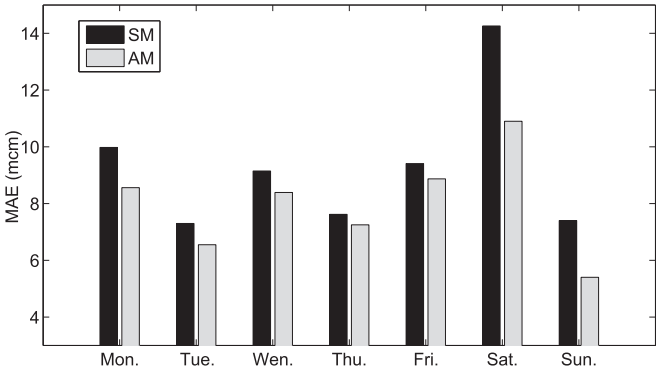

Fig. 13. Comparison of average prediction MAE for every day of the week between the SM and the AM by using the FNF-SVRLP method.

Acknowledgements

This work was supported by National Grid, UK. The authors wish to thank National Grid for providing useful data for our research, and valuable discussion on the results obtained.

References

[1] Parikh J, Purohit P, Maitra P. Demand projections of petroleum products and natural gas in India. Energy Oct 2007;32(10):1825–37.
[2] Wadud Z, Dey HS, Kabir MA, Khan SI. Modeling and forecasting natural gas demand in Bangladesh. Energy Policy 2011;39:7273–81.
[3] Soldo B. Forecasting natural gas consumption. Appl Energy 2012;92:26–37.
[4] Almeida AT, Lopes AC, Carvalho A, Mariano J, Jahnd A, Broeged M. Examining the potential of natural gas demand-side measures to benefit customers, the distribution utility, and the environment: two case studies from Europe. Energy 2004;29:979–1000.
[5] How electricity is generated in the UK. http://www.hi-energy.org.uk/Renewables/Why-Renewable-%20Energy/How-electricity-is-generated-in-the-UK.htm [accessed 07.06.2014].
[6] Electricity generation and supply figures for Scotland, Wales, Northern Ireland and England. 2008 to 2011 [accessed 07.06.2014], https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/65841/7345-elec-gen-2008-2011-et- article.pdf/.
[7] Sarak H, Satman A. The degree-day method to estimate the residential heating natural gas consumption in Turkey: a case study. Energy 2003;28:929–39.
[8] Siemek J, Nagy S, Rychlicki S. Estimation of natural-gas consumption in Poland based on the logistic-curve interpretation. Appl Energy 2003;75(1–2):1–7.
[9] Gutierrez R, Nafidi A, Sanchez RG. Forecasting total natural-gas consumption in Spain by using the stochastic Gompertz innovation diffusion model. Appl Energy 2005;80(2):115–24.
[10] Potočnik P, Thaler M, Govekar E, Grabec I, Poredos A. Forecasting risks of natural gas consumption in Slovenia. Energy Policy 2007;35:4271–82.
[11] Erdogdu E. Natural gas demand in Turkey. Appl Energy 2010;87(1):211–9.
[12] Brown RH, Kharouf P, Feng X, Piessens LP, Nestor D. Development of feed-forward network models to predict gas consumption. In: IEEE international conference on neural networks C conference proceedings, vol. 2; 1994. p. 802–5.
[13] Brown RH, Matin I. Development of artificial neural network models to predict daily gas consumption. In: IECON proceedings (industrial electronics conference), vol. 2; 1995. p. 1389–94.
[14] Khotanzad A, Elragal H. Natural gas load forecasting with combination of adaptive neural networks. In: Proceedings of the international joint conference on neural networks, vol. 6; 1999. p. 4069–72.
[15] Khotanzad A, Elragal H, Lu TL. Combination of artificial neural-network forecasters for prediction of natural gas consumption. IEEE Trans Neural Netw 2000;11(2):464–73.
[16] Ivezić D. Short-term natural gas consumption forecast. FME Trans 2006;34:165–9.
[17] Azadeh A, Asadzadeh SM, Ghanbari A. An adaptive network-based fuzzy inference system for short-term natural gas demand estimation: uncertain and complex environments. Energy Policy 2010;38(3):1529–36.
[18] Salahshoor K, Kordestani M, Khoshro MS. Fault detection and diagnosis of an industrial steam turbine using fusion of SVM (support vector machine) and ANFIS (adaptive neuro-fuzzy inference system) classifiers. Energy 2010;35(12):5472–82.
[19] Zhang WY, Hong WC, Dong Y, Tsai G, Sung JT, Fan GF. Application of SVR with chaotic GASA algorithm in cyclic electric load forecasting. Energy 2012;45:850–8.
[20] Taghavifar H, Mardani A. A comparative trend in forecasting ability of artificial neural networks and regressive support vector machine methodologies for energy dissipation modeling of off-road vehicles. Energy 2014;66:569–76.
[21] Hong WC. Load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm. Energy 2011;36:5568–78.
[22] Lau KW, Wu QH. Local prediction of non-linear time series using support vector regression. Pattern Recognit 2008;41(5):1556–64.
[23] Gong XF, Lai CH. Improvement of local prediction of chaotic time series. Phys Rev E 1999;60(5):5463–8.
[24] Packard NH, Cratchfield JP, Farmer JD, Shaw RS. Geomertry from a time series. Phys Rev E 1980;45(9):712–26.
[25] Takens F. Detecting strange attractors in turbulence. Lecture Notes in Mathematics, vol. 898. Springer; 1981. p. 366–81.
[26] Sauer T, Yorke J, Csdagli M. Embedology. J Stat Phys 1991;65(3/4):579–616.
[27] Grassberger P, Procaccia I. Estimation of the kolmogorov entropy from a chaotic signal. Phys Rev A 1983;28:2591–3.
[28] Liebert W, Schuster HG. Proper choice of the time delay for the analysis of chaotic time series. Phys Lett 1989;142:107–11.
[29] Data Item Explorer. http://marketinformation.natgrid.co.uk/gas/ [accessed 02/05/2014].
[30] Kusiak A, Zheng H, Song Z. Short-term prediction of wind farm power: a data mining approach. IEEE Trans Energy Convers 2009;24(1):125–36.
[31] Lora AT, Riquelme JC, Ramos JLM. Influence of kNN based load forecasting errors on optimal energy production. Lect Notes Comput Sci (Springer Berlin) 2003;2902:189–203.
[32] Sorjamaa A, Reyhani N, Lendasse A. Input and structure selection for k-NN approximator. Lect Notes Comput Sci (Springer Berlin) 2005;3512:985–92.
[33] Elattar EE, Goulermas JY, Wu QH. Generalized locally weighted GMDH for short term load forecasting. IEEE Trans Syst Man Cybern Part C: Appl Rev 2012;42(3):345–56.